Clutter And Realism

D.Mishev
B.Salters
M.Wijntjes
J.Verlinden

PHILIPS

Delft
University
of Technology

Clutter and Realism

An investigation on the effect of clutter on the perceived realism of 3d computer generated renderings.

Industrial Design and Engineering Technical Report (April 2012)

Deian Mishev; Bart Salters; Maarten Wijntjes; Jouke Verlinden;

Technical University of Delft , Philips Research Eindhoven

This dissertation is submitted to the faculty of the Technical University of Delft the Netherlands and the Visual Experience Department of Philips Research Eindhoven. This dissertation is a partial fulfillment of the requirements for the degree of Master of Industrial Design and Engineering from the faculty of Industrial Design of TU Delft.

Acknowledgment:

First and foremost I would like to thank my friends and colleges from the department of Visual Experience in Philips for being actively and enthusiastically engaged into perceptual research in the field of virtual reality and virtual prototyping. Their help and guidance has been important for this project. I would also like to personally thank Mariska Stokkermans , Michael Murdoch and Dragan Sekulovski for taking part in the initiation and facilitation of this project in the department. Besides them I would also like to thank the entire board of the Philips Internship Committee which gave me a lot of moral support along the way and some priceless memories from my free time in the city of Eindhoven. Last and not least I would also like to thank my mentors and all professionals involved in the fields evolving around this work. To me their hard work and efforts are a constant professional motivator as well as a source of inspiration.

Content:

1. Abstract 2
2. Introduction 3
3. Clutter 3
4. Previous Work 4
5. Current Project 5
6. Technical Set Up 5
    6.1. Capturing Technology 5
    6.2. 3d Modeling Technology 6
    6.3. Rendering Technology 7
    6.4. Tone Mapping software and Key Operator 7
    6.5. Display Technology 8
7. First Experiment 8
    7.1. The Scene 8
    7.2. Variables Introduction 9
    7.3. The Test Methodology 9
    7.4. The Scale 10
    7.5. First Experiment Conclusions 11
8. Second Experiment 13
    8.1. The Scene 13
    8.2. Variables Introduction 13
        8.2.1. Angle Shift 13
        8.2.2. Displacement 15
    8.3. Test Methodology 16
    8.4. Second Experiment Conclusions 16
9. Third Experiment 21
    9.1. Test Methodology 21
    9.2. Third Experiment Conclusions 21
10. Project Conclusions 24
11. Recommendations 25
    11.1. Specialized Applications 25
    11.2. Commercial Applications 25

References 26
Appendice 1. First test: Objects Introduction
Appendice 2. Paper Print Questionnaire

April 2012

Clutter & Realism

An investigation on the effect of clutter on the perceived realism of 3d computer generated renderings.

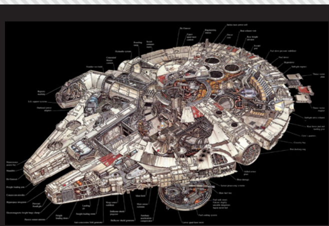
Technical University of Delft & Philips Research Eindhoven

Keywords: Clutter, realism, computer graphics, virtual prototyping, 3d objects;

1. Abstract

Current Philips practices include development of physical environments in which luminaires can be tested with the user. However Philips is interested in investigating how well virtual prototypes can be used to predict the outcome of non-virtual tests. For this, Philips is turning towards identifying a number of features in 3d environments which influence the human perception of realism. This study utilizes traditional psychophysical analysis to investigate the effect of object induced clutter and its connection to the perception of visual realism in 3d renderings. This project goes beyond the identification of pre-perceived common trends and reveals the actual perceptual value of clutter according to the perceived realism that it brings to virtual scenes.

Fig.1

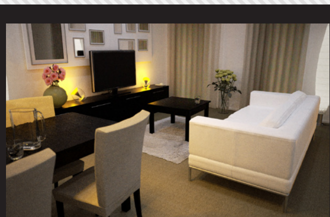


Millennium Falcon
Star Wars Episode IV: A New Hope

A spacecraft commanded by smuggler Han Solo (Harrison Ford) and his Wookiee first mate. Drawing built on models.

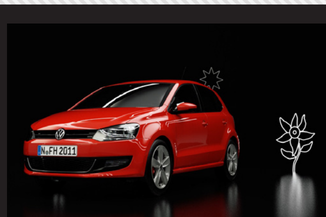
Water Glasses and a Jug
D.Mishev - 2012

A test scene created to visualize the capabilities of the Indigo rendering engine.



Philips Experience Laboratory
D.Mishev - 2012

A renderer created by request of the visual perception cluster in order to test different light set-ups.



VW Polo © christoph.kohlhas
Christoph Kohlhas

3ds Max and Glare Technologies "Indigo Renderer" webpage demonstration rendering.

Autor
D.St.Mishev (Deian)
Integrated Product Design
Technical University of Delft

Company Supervisor
Ir. B. Salters (Bart)
Philips Senior Scientist
Visual Experience Group

University Mentor
Dr. M. W. A. Wijntjes (Maarten)
"ID - HICD" Human Information Communication Design

University Supervisor
Ir. J.C. Verlinden (Jouke)
"DE – CADE "Computer Aided Design & Engineering

2. Introduction

Real life luminaire prototyping obviously lacks the flexibility and functionality of its virtual counterpart. It is easy to see why you would have use of software allowing you to change material properties and light intensities without having to recreate prototypes and spend a lot of resources during the development of the luminaires.

However in the field of virtual luminaire prototyping there is also a ever present level of quality that should be met for all of the features of the virtual environments. “Light” or “visible light” is the electromagnetic radiation that is visible to the human eye and is responsible for the sense of sight **(1)**. Simulation of light bounces is essentially what the rendering engines do in order to produce the visualizations. Thus to simulate light which is physically correct and expect it to be visually comparable with the illumination produced by a real luminaire you have to take special care to address all of the features of the virtual environment including surface properties, material maps, number of light bounces, etc. The current developments in 3d rendering technologies (ex: Maxwell, Indigo, Mitsuba, Lux, Octane…) incorporate advanced algorithms simulating physically correct light bounces and real life physical features of both objects and environments (no approximation rendering). These rendering packages allow designers to create visualizations enabling their users to experience the end products on a level far beyond the conceptual.

The current techniques for pursuing visual realism in graphics are subdivided into different categories: geometric modeling, rendering, behavior and interaction **(2)**. Currently, for all of these, the methods and algorithms used to achieve realism are dependent on both the application of the renderings, the type of user and the type of resources. For example, when looking at a map some may not care about the fine detail while other will want all of it. Then again, more details definitely take more time while the perceptual value they bring for the sensation of realism stays largely undefined. As Philips is interested in creating visualizations to target none experts in the field the company has a drive for innovating, in not only the ongoing rapid development of technologies but also, in the development of a deeper understanding of what the prominent features of virtual environments are and how to manage them accurately.

3. Clutter

Clutter as argued by CG artists such as Bill Fleming is one of the most obvious traits of visual reality **(3)**. Still in virtual environments it can take on many forms, and largely depends on the tasks being performed with the visualizations. A example of that is the development of the “grunge pass” (a render layer adding material clutter as dirt and stains) made by George Lucas during the first Star Wars. George Lucas wanted all of the space hardware featured in his movie to look used and he gave it more texture which simulates dirt and other surface imperfections. Before that, model makers would rarely think of dirtying up their meticulously crafted models. Then, visual effects had a look which now might be considered “too clean”. Still the word “clutter” is generally one used to address the scene in reference to the visual effect that the objects in it excite. This more corresponds to the dictionary definition of clutter which is **(4)**:

Noun 1. a disordered heap or mass of objects
2. a state of disorder

This is roughly the type of clutter that is addressed in this project. Nevertheless the dictionary definition needs to be reformulated in order to suit this study in virtual environments. Thus the following work definition is embraced:

“Clutter is the number of objects present, relative to the amount of display space available, and their placement in that space relative to the messiness or orderliness of it.”

This definition is embraced as according to the designer it summarizes the key features of clutter accurately. Thus the designers assumption is that the clutter level of the scenes is affected by the number of visible objects in the scene and their position. This assumption of the designer is one which this project does not attempt to confirm. Therefore the definition of clutter adopted is left constant and no attempt to investigate what different participant understand by "clutter" is made.

Still reducing the number and detail or objects in the scenes increases the productivity of the 3d hardware system under any conditions and scenarios. This is by decreasing the overall workload of these systems and specifically the one on graphic bottleneck stages such as the rendering and live viewing. In order to better control the realism of virtual scenes the effects of different amounts of clutter needs to be investigated. Thus, the research question arises:

"What is the effect of the amount of clutter on the perceived realism of 3d renderings?"

The goal of this project is to:

"Measure the effect of clutter on the perceived realism of virtual environments"

This is done in order to control the balance between the amount of resources (modeling time, rendering time, system hardware, etc.) and the visual perception of realism.

4. Previous Work

There are studies conducted in the field of virtual reality which show that the visual realism of a virtual environment can significantly improve a users experience and indeed increase his/her sense of presence in that environment (5). This in fact can easily be referenced to the perception of visual realism that we experience. In work previously done in Philips "A comparison of perceived lighting characteristics in simulations versus real-life setup" (6), Salters and Seuntiens conclude that:

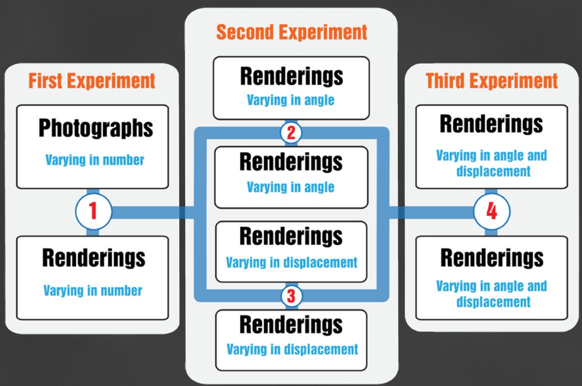
"Although absolute values differ between real life experiments, and simulations, relative values follow the same trends"

According to their work this implies that a better match in absolute numbers should be possible. If so, this will greatly speed up design and evaluation of future luminaire designs. The context and framework for this project is dominated by this need of Philips Electronics. Subsequent to the above stated project, the company has developed an interest in investigating how well a virtual prototype can be used to predict the outcome of a non-virtual test.

Still in research in the field there have been several different approaches to the creation of realistic images. One is to analyze, measure, approximate, and simulate the various physical processes that form a real-world image (models, light transport, surface BRDFs (7), and more). However the few projects that have created and compared 3d images to specific target photographs (ground truth) rarely reveal which visual factors a viewer expects in order to perceive the images as real or unreal (8),(9). Another common approach to simulating visual realism is image-based rendering. It's known to created synthetic images that are nearly indistinguishable from photographs. Unlike geometry based rendering, image-based rendering works by rearranging image samples taken directly from photographs (9) and projecting these on 3d surfaces which are again either generated or modeled. There is a great probability that these have a constant level of realism when compared with photography. Nevertheless most current methods for generating image based renderings can still only be sufficient in simulating a few of the aspects of virtual environments which are important to prototyping of luminaires. This is because the environments generated by these methods lack features of geometry and materials which are very important to how light is perceived in larger scenes or the perceptual quality of their outcomes has not been scientifically validated. Most closely related previous research can more easily be compared to virtual product prototyping and interior design then the prototyping of luminaires. Sadly this is mostly insufficient and makes it hard to find true references between this project and the ones previously done in the field of realism and how it is affected by object features of its virtual environment.

5. Current Project

This project presents a more comprehensive approach to scaling clutter in scenes which are a lot more complex in their lighting, models, material textures, etc. Thus



Project stages Fig.2

the techniques utilized for this project are ones which Philips wants to adopt in its future luminaire research and development cycles while still striving to effectively link to the previous work done both inside and outside of the company.

For the first experiment of this project we compare renderings and Tone Mapped (10) HDR photographs and the effect that the number of objects in them has on their visual realism. This defines a set assumption that high fidelity photographic images are a close physical representations of a real still scene. For the following stage renderings are compared between each other. These according to the objects they contain are varying in angle and placement. This project is comprised of 3 experiments made of a total of 4 participant test stages (Fig.2). This project strives to define and manage clutter and its affect on the sensation of visual realism in renderings. As the feature in focus is the object induced visual clutter the project begins with the hypothesis that:

"The amount of clutter of objects strongly affects the perception of realism in 3d renderings."

During the development of this project the above stated is thoroughly tested and the needed conclusions are drown according to the findings that are made. The design of the series of tests in the perceptual experiments is governed by the need to create the large sets of images scaling the clutter features as accurately as intended. These sets are comprised of photographs and renderings or only renderings. The software and hardware used to create these images are chosen to most accurately serve the goals of this experiment and create a project environment in which the participants can better focus on the features of the experiments. The technology used to prepare and administer this tests is comprised of capturing, 3d modeling, rendering, tone mapping, color calibration and display technology.

6. Technical Set Up

6.1. Capturing Technology

The canon 450D

The Canon SLR digital camera has been chosen as it is commonly used amongst both professionals and amateurs. With it, we capture 2 different sets of images for the first stage of the experiment and all of the texture maps needed. The experiment photography were done with a 28 mm lens which was also recreated for the cameras in the virtual scenes. The 2 sets are in general the same scene shot from two different camera views. These are made with three different exposure times (1/20, 1/80, 1/5 s.). These shots are visualy ranging from:

Scene space photography Fig.3

Too dark – where parts of the image are clipped in black.
Normal – the exposure is neither clipped in black nor clipped in white.
Too bright – where parts of the image are clipped in white.

Essentially that is done in order to recreate the process in which the renderings are made. The photographs are then combined into a HDR image with HDR software and afterwards scaled down to an LDR image with the same key features and TMO (tone mapping operator) as the renderings. These are cropped to the size of the renderings (Fig.3) (1920x1280) and color calibrated to match the output specifics of the display technology used.

6.2. 3d Modeling Technology

Autodesk 3D Studio Max

Autodesk 3ds Max, is a modeling, animation and rendering package developed by Autodesk Media and Entertainment. It has modeling capabilities, a flexible plug-in architecture and can be used on the Microsoft Windows OS. It's frequently used by video game developers, TV commercial studios and architectural visualization studios. It is also used for movie effects and movie pre-visualization. In addition to its modeling and animation tools, the latest versions of 3ds Max also features shaders (such as ambient occlusion and subsurface scattering), dynamic simulation, particle systems, radiosity, normal map creation and rendering, global illumination, a customizable user interface, and its own scripting language (11).

3d Studio Max was the choice for a modeling tool as it allows a relatively easy manipulation of multiple objects, can work with proxies and is regularly used by the interior design community which can be referred for help in forums. In addition to the modeling capabilities of 3d Studio Max it is one of the modeling tools which has good integration with most rendering engines and their exporters.

6.3. Rendering Technology

Indigo renderer (sample image Fig. 4)

Indigo renderer sample image Fig.4

Indigo Renderer is a 3D rendering software that uses unbiased rendering techniques to create photo-realistic images. In doing so, Indigo uses equations that simulate the behavior of light, with no approximations or guesses taken. By accurately simulating all the interactions of light, Indigo is capable of producing:

- Depth of field, as when a camera is focused on one object and the background is blurred.
- Spectral effects, as when a beam of light goes through a prism and a rainbow of colors is produced.
- Refraction, as when light enters a pool of water and the objects in the pool seem to be "bent".
- Caustics, as in light that has been focused through a magnifying glass and has made a pattern of brightness on a surface.

Indigo uses methods such as spectral light calculus, and virtual camera model. Scene data is stored in XML or IGS format. Indigo features Monte-Carlo path tracing, experimental support for bidirectional path tracing and MLT (Metropolis Light Transport), distributed render capabilities, and progressive rendering (image gradually becomes less noisy as rendering progresses). Indigo also supports subsurface scattering and has its own image format (.igs). Indigo renderer 3.0 includes the Autodesk 3DS Max 2012 compatibility through its 3ds plug-in named Maxigo. Maxigo is a *.igs file exporting plug-in which allows for file transfer between the formats of the two programs. The combination of this exporter and the 3D Max software allows for the use of IES files specifying different light intensities and directions. The Indigo renderer online databases also include IES files which store realistic properties of different light emitters. Including all of the above Indigo is also chosen because of the free access in the Philips Visual Experience Department and its no approximation characteristics. These combined with the use of the photometric data files for the light set ups are claimed to generate images that are comparable to photographs (12), (13).

6.4. Tone Mapping software and Key Operator

Luminance HDR and Reinhard tone mapping model

Luminance HDR (14) is an open source image editing application that aims to provide a workflow for generating and editing high dynamic range (HDR) and low dynamic range (LDR) images. Luminance HDR is used as a tone mapping tool because of its support for workflow with the most popular High and Low dynamic range formats. Tone mapping is a technique used in image processing and computer graphics to map one set of colors to another in order to approximate the appearance of high dynamic range images in a medium that has a more limited dynamic range. In the case of this project the medium is the display technology described in the following point.

Luminance HDR is used with the available Reinhard tone mapping model. Reinhard is a tone mapping model converting high dynamic range images to low dynamic range images. Although the goals of tone mapping can be differently stated depending on the particular application Philips has so far found the Reinhard Tone mapping method suited for applications where creating subjectively satisfactory and essentially artifact-free images is the desired goal (15).

6.5. Display Technology

When looking for a convenient display technology both the brightness and the Delta-E (ΔE) index of the display need to be considered. Brightness is the perception elicited by the luminance of a visual target. For display technology it is usually expressed in Candela per square meter (cd\m^2). Delta E is a measurement used to indicate how much a color deviates from an accepted standard. The higher the ΔE, the more inaccurate the color. Perfect color has a ΔE of zero. However, we need not achieve a ΔE of zero. This is because the human eye is only capable of detecting color difference at certain thresholds. The minimal detectable difference is about 1 ΔE. Calibration generally seeks to achieve a ΔE for white and all of the primary and secondary colors of no more than 2.5. However, unless a display has a fully-realized color management system (the vast majority do not), it is unlikely that calibration can achieve this level of performance for the primary and secondary colors. On the other hand, since virtually all displays have a full complement of gray scale adjustments, calibration can usually meet this standard for white **(16)**.

For the tests a 42" Philips display is chosen. The display is a full HD display with a maximum luminance of 230 cd/m2. When non color calibrated images are used on it the ΔE measures at about 5.91. However for the tests all images are color calibrated and reach a ΔE level of 1.81. This is sufficient for the purposes of this experiment.

7. First Experiment

The first experiment is a quantitative approach towards discovering what affects peoples' perception of realism in images varying in clutter. Clutter is represented by only a change in the number of objects present in the scene. The choice of where to place these objects is made by the designer. Thus besides his personal preference the placement is only governed by physics as obviously some objects in the scene are supported by others and would be left hanging in mid air if introduced first. The images in this test are both photographs and 3d renderings. The photographs are made in an office space named "Light Laboratory". This laboratory is regularly used for perception tests with light. For the required images however, it is used for only its spacial characteristics and the fact that it provides for an environment which can be freely manipulated for the set up of the photographs **(Fig.3)**.

7.1. The Scene

The scene is chosen to stand as a typical office room, which has no defined usage. The randomly placed objects in it are made according to the judgment of the designer. In the renderings these objects are models of items found inside the "Light Laboratory" or around the office spaces of the Visual Experience Group in Philips. According to the judgment of the designer these items are supposed to look as if they've been used but afterwards left randomly about. While this might not be that uncommon for an office space, the scene (again according to the opinion of the designer) has no clear set up. Nevertheless some areas in the space will hopefully suggest to the viewer that at a point office work has been performed in them. That was an additional intention of the designer. The desire was to create a room featuring a set of objects which were placed in a way that compels the viewer to look around it and try to examine it to the fullest. The objects are introduced in an order which (by the opinion of the designer) balances the scene visually.

The space is labeled "Storage Common Room". The subjects are not given this definition or any other information about the space. This aims to minimize any factor which might affect their unbiased judgment of it.

7.2. Variables Introduction

For the first user test, by "objects" we mean any that are considered not to be architectural elements such as doors, windows, flooring etc. This changes in the second stages but is addressed later. The objects in the scene, despite varying in number, are models of common items, which can be found in the actual environment. In the 3d scene the images needed for the material maps are also directly taken from the actual objects.

For this test we disregard the positioning and orientation of the objects (except of the issues stated in point 7 of this report). The number of objects ranges from 6 to 30 in a step increase of 6 (6(Fig.5),12,18,24,30(Fig.6)). The order and the list of the object introduction in the scene are given in the appendice (A1).

Fig.5

The above image is the low(6 object) level.

Fig.6

The above image is the high(30 object) level.

The images which are used in the first experiment are both photographic and 3d generated. The 10 photographic images are shot from two different camera angles used to both fool the user about the genuinity of all images and present the contradiction to renderings which is the challenge for peoples' perception of realism. These are both represented by two camera angles meaning that the overall number of scenes scales to 20 (5 levels * 2 views * 2 types (photographs and renderings)) (Fig.7). The renderings and photographs are set to resemble each other by all features.

7.3. The Test Methodology

12 users were faced with evaluating the 2 sets of 40 images (the 20 images randomly shown twice) on a 7 point Thurstone scale **(17)**. These images are tone mapped, color calibrated and presented on a 42 inch flat screen TV in a darkened room with a constant light intensity which is maintained throughout all the stages of the experiment. The subjects are situated about 1,2 meters away from the display. For the this experiment each image is viewed as long as the participants required to make their choice. Amongst all there are 2 pilot users which are added to the main number as they represented no significant change to the test methodology. All of the participants are not told which images are photographic or 3d generated. Instead they are asked to evaluate the realism of the images for the first show set of 40 and the level of clutter for the second show set of 40.

The task for the first show set of 40 is:

Evaluate the overall realism of the presented scene on the seven point scale.

The task for the second show set of 40 is:

Evaluate the overall level of clutter of the presented scene on the seven point scale.

In the case that the subjects asked for a the definition of realism, they were encouraged to interpret it on the basis of whatever preconceived notion they have for it. For the second show set of clutter however the users were given the definition adopted by this project.

7.4. The Scale

The seven point scale ranges from 1 (absolutely unrealistic) to 7 (absolutely realistic). The seven point scale is chosen because globally, it is apparent that it shows better properties than the 5-point one (17). It has been proven that higher point scales generally show higher validity than the 5-point scales. Coelho and Esteves find the results of 5 point scales as more inaccurate. This confirms a hypothesis that within an odd scale the middle point is often used by the respondents that prefer to reduce the response effort, resulting in series of "evening out" scores. This is something that is highly undesirable for the current experiment and still the seven points scale used also has a middle value. This is as most previous research in the field, that were done in Philips, have been administered on it. Therefore it is an obvious choice before the 10 point scale for the reason that the data derived can be compared with the one that Philips Research already has. Still as pointed out by Coelho and Esteves, people can deal with a higher number scale than the 5 point Likert scale and these scales usually score more definitive choices.

The scale method is an "equal-appearing intervals" method (18). This method requires the observer to place samples in categories of equal-appearing intervals. Equal-appearing is equivalent to saying the categories are of equal width, except of course for the end categories. Categories are labeled with names like "good," "better," "best," or with numbers and some

Fig.7

Storage Common Room

On the left you can observe the two renderings and on the right of the two photographs. They both are tone mapped and colour calibrated with the exact same methodology. In addition sampling images from the physical scene were used to create the image maps assigned to the materials of the objects in the renderings. Both types of images appear green as they were colour calibrated to fit a specific screen instead of print.

ranked numbering. As with ordinal scaling, this method of data collection does not directly yield an interval scale without additional assumptions or models. Average observers typically distinguish about 7 different categories **(19)**, so additional categories may contribute very little "scale information". Many categories require the observer to make fine intervals of discrimination amongst them, while a smaller number does not. Still we presume our participants to be none experts but expect them to be familiar or interested in the field as they are all Philips researchers, employees, interns, etc. The 7 point scale chosen is digitally administered during the viewing of each image and when the choice is made the script commands a random change of images until all are seen twice.

7.5. First Experiment Conclusions

The conclusions of the first user test (as administered according to the described methodology) are presented starting with clutter and followed by realism (Fig.8 and 9). The bars above each of the measurements represent the standard error for each of the mean answers. In the second set of the first test (Fig.8) users scaled clutter according to the amount of objects in the scene. The scale thus shows a rise with the rise of number of objects in the given images. The standard deviation error bars of the clutter test shows that there were generally less fluctuating answers for the photographs but nevertheless it's clear that the users respectfully put all images in a scaling order which is in a progressive relation to the number of objects. With the scores of realism (Fig.9) however, users demonstrated that they do not consider the rise of object number to be directly linked to a rise of the realism of the scenes. This is despite that a definition of realism was not given and the stage asking about it was administered first. Photographs scored consistently higher on their level of realism. Both photographs and renderings have error bars which show that the scale answers differed greatly in between the users. Nevertheless, after a further look inside the data with SPSS (IBM SPSS Data Collection) we were able to divide the subjects into groups which demonstrate similarities in their attitude towards the realism of the presented scenes (Fig.10)(Fig.11).

Fig.10

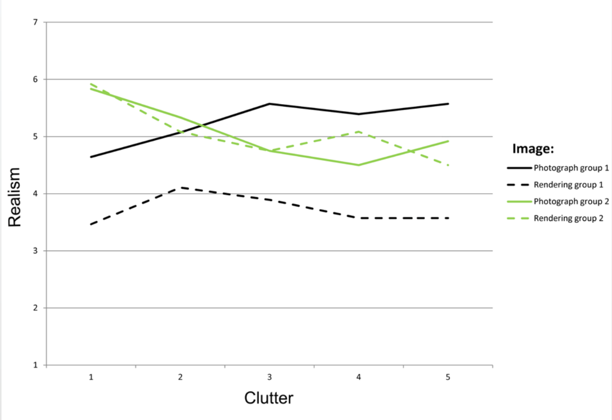
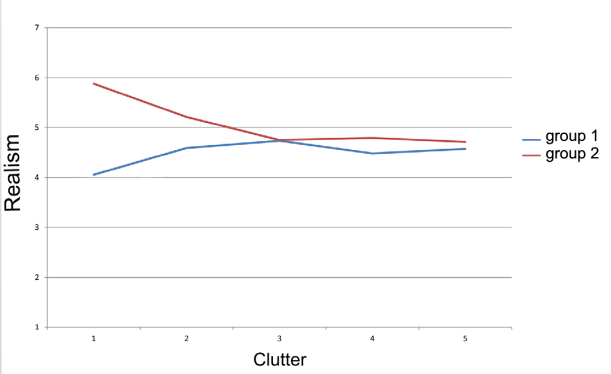


Fig.11



Group one is by far the largest group in this analysis. It's comprised of participants number: 1,2,4,5,9,10,11. Definitive for this group is the fact that they set the gap between the photographs and renderings (Fig.10). Group two is comprised of participants number: 3,8,12. These are the participants which considered that there is much less difference in realism between photographs and renderings. Their number is lower than the people who create the gap between the same. These participants however scored both photographs and renderings as generally scaling down in realism, as the number of objects increases (Fig.11). Participants number 6 and 7 are considered to be outliers in this cluster analysis and thus are not added to the graphs above. Nevertheless, this interesting additional SPSS clustering is only noted and not pursued any further. It is however important to mention that the participants of this experiment were all in their mid twenties and differed the most by their nationality than anything else.

As a general conclusion of this stage the decision is made to focus on the other features of clutter for the second experiment in this project. These are angle shift and displacement which are more correspond to the messiness and orderliness of clutter than the sheer number of objects in the scene.

Fig.8

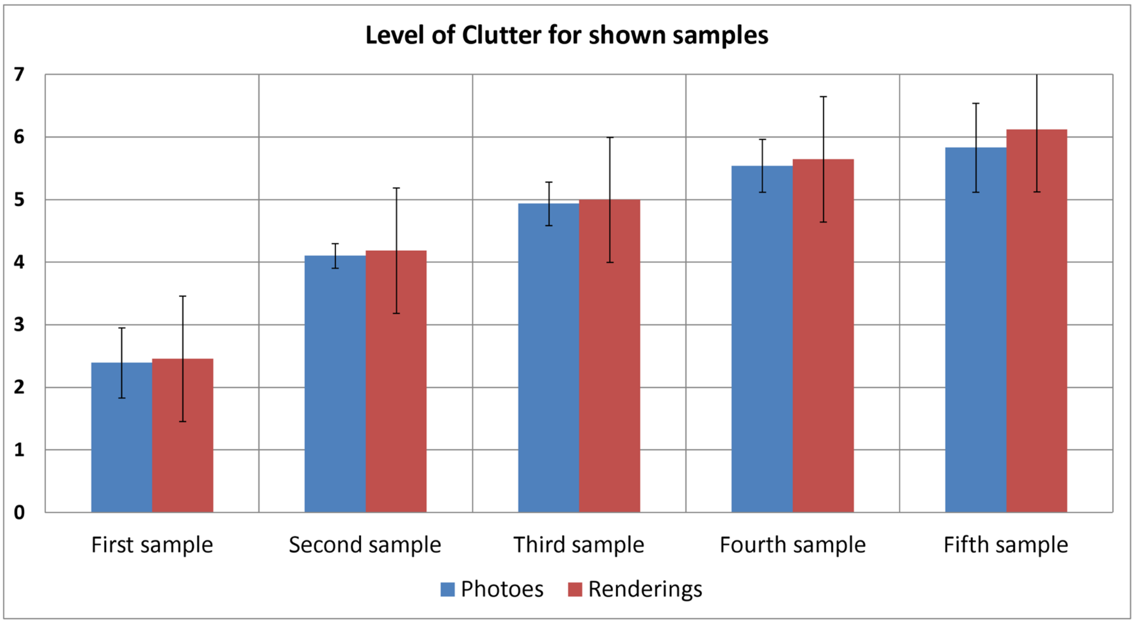
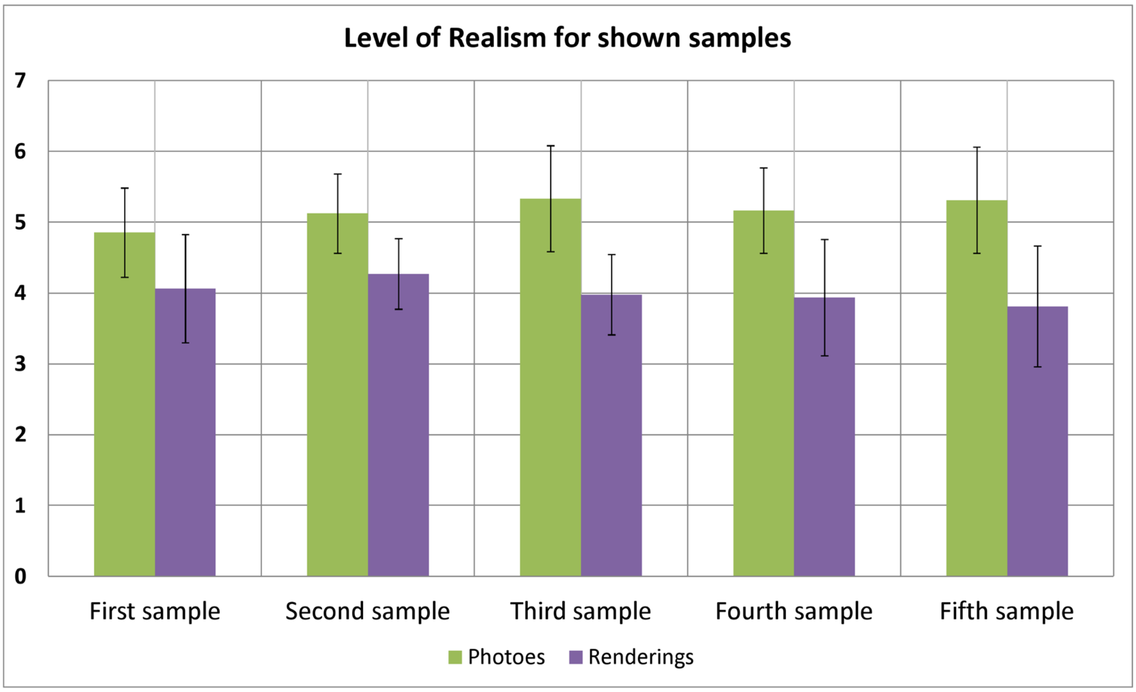


Fig.9



8. Second Experiment

8.1. The Scene

The images made for this scene are modeled and rendered to present a similar office environment as the environment from the first experiment. The scene is chosen to be an office table. Therefore in this instance the table is considered to be the equivalent of what the architectural elements were in the first scene. Sixteen common office objects are dispersed on the table.

8.2. Variables Introduction

The variables of the second set of images are more definitive for the features of clutter which are associated with its chaos feature than the sheer number of objects in the environment. The angle and placements shift of the objects are in this part varied in a way which is random and yet quantifiable (Fig. 12).

8.2.1. Angle Shift

For the first stage of the second user tests the orientation (as a scale of clutter) of objects and its effect on the perception of realism is tested. For this stage there are no photographs. This is as during the first stage, renderings showed that they were perceived as less real but with a constant offset from the level of realism of the photographs. Thus for the second stage the created renderings are evaluated in between each other.

Sixteen common office objects are dispersed in the scene. These are again placed according to the designers personal presumptions of what placement the objects

Fig.12

	Object name:	Angle 180	Angle +/- 5	Angle +/- 10
1	Black Mug	180	171.68	171.00
2	Yellow Mug	180	180.08	169.79
3	Disc above	180	176.32	172.28
4	Disc bellow	180	183.61	174.71
5	Keyboard	180	174.96	172.65
7	Sheet 1	180	179.39	182.06
8	Sheet 2	180	180.52	177.01
9	Pen 1 Blue	180	180.48	177.57
10	Pen 2 Green	180	185.38	178.97
11	Pen 3 Brown	180	185.73	187.28
12	Remote Control	180	178.10	159.08
13	Monitor	180	171.25	177.31
14	Marker 1	180	185.98	174.66
15	Marker 2	180	180.40	181.22
16	Marker 3	180	182.71	186.58

The generated angle shift is done by a randomness formula which introduces a predefined mean standard deviation shift from 180 (base state) and randomness multiplication coefficient (RAND()) which the Microsoft Excel generates to be any in between 0 and 1. In the two cases the formulas are:

= **NORMINV (RAND(), 180, 0)** – for the resulting base state
= **NORMINV (RAND(), 180, 5)** – for the 5 degree standard deviation shift
= **NORMINV (RAND(), 180, 10)** – for the 10 degree standard deviation shift

0 deg. of angle shift

Fig.13

5 deg. of angle shift

Fig.14

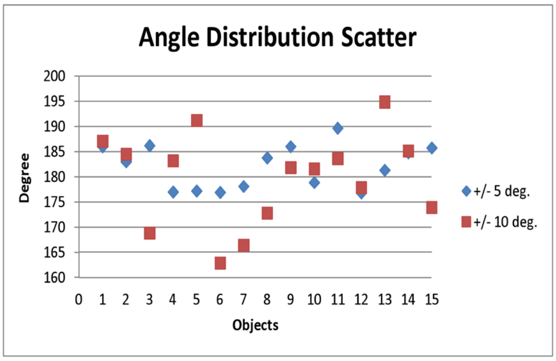
10 deg. of angle shift

Fig.15

could have. Thus the objects only scale in their angle orientation to a predefined world scale. This orientation is according to the vertical axis of the 3d models. This is again because the research is intended to stay within the borders of what's physically possible. Thus any abnormalities such as objects suspended into space or lifted with one side from the surface of the table are undesirable. This shift in angle defines the variable tested as "angle shift" (Fig. 13, Fig. 14, Fig. 15).

Through the formula described in (Fig.12) we are producing this angle shift for each of the objects. After the resulting shifts for 5 and 10 degrees are applied the scenes are rendered from 5 different views. The graph with the resulting scatter plot (Fig.16) for the 5 and 10 degree below, demonstrates the effectiveness of the method used to scale the angle shift:

Fig.16



The views are taken from virtual cameras placed on the surface of a sphere which has its center in the center of the table. This ensures that there is a constant distance in between the camera and the objects. The necessity for this comes from the conclusions of a pre-test questionnaire administered in paper format (A2). This was focused on defining some of the "none clutter" and "none technical" features of the scenes which were tested. From the comments after the first research, the questionnaire pointed towards focusing on the details of the objects, their maps and surfaces. The questionnaire also lead to keeping objects in the virtual space at a controlled distance from the virtual cameras. This was done by placing the 5 cameras on the surface of a sphere having its center in the center of the viewing space. This ensures that the objects would have a relatively constant distance from the camera which showed as necessary for maintaining a constant level of realism in between camera views (demonstrated in A2).

The paired comparison was administered digitally under the exact same settings as the settings from the first test. The participants were asked to pair compare each of the sets choosing the one they find as more realistic. They were informed about the overall number of comparisons (60) and told that they have a total of 30 minutes to complete the test (30 seconds for an image). The script made shows the chosen paired sets in a random order until they are all compared a predefined number of times.

8.2.2. Displacement

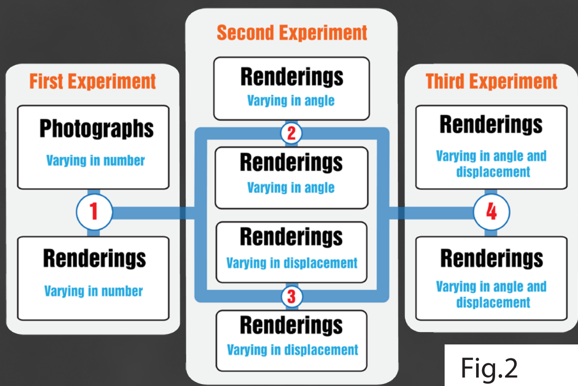
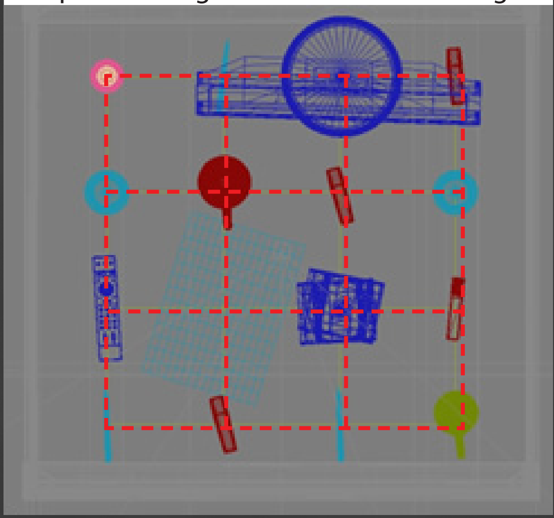


Fig.2

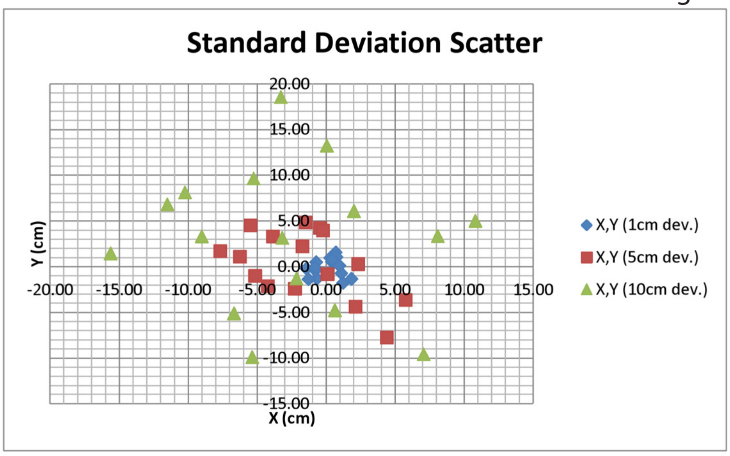
For the second stage of the second experiment the placement as a scale of clutter of objects and its connection to the perception of realism is tested (Number 3 in Fig.2). For this stage there are also no photographs and thus renderings are tested in between each other. The images made are modeled and rendered to present a similar office environment as the environment from the previous stages. The scene is the same office table as the previous stage. Therefore in this instance again, the table is considered to be the equivalent of what the architectural elements were in the first scene. The 16 common office objects are now evenly aligned within a 4x4 grid (Fig.17). The angle

4x4 placement grid Fig.17



orientation chosen is again according to the vertical axis of the 3d model and is the exact same angle shift as the one the objects had for the 10 degree standard deviation. However, one scene sample with 5 degree standard deviation is included and compared with the rest. Again the same formula is chosen to randomize the position. Here we've been careful to keep all objects within the table. The cm. of standard displacement have been selected to be: 0 cm, 1 cm, 5 cm, 10 cm (all with the 10 degree of standard deviation shift) and one 5 cm and 5 degree standard deviation sample. The resulting scatter plot (Fig.18) for the 1, 5 and 10 cm of angle shift demonstrates the effectiveness:

Fig.18



The set of images on the following page demonstrate the top left view of the 0, 5 and 10 cm. (top down) of standard deviation displacement (Fig.19, 20 and 21).

8.3. Test Methodology

All samples are pair compared in between each other in respect to their view. The paired comparison is again administered digitally under the exact same settings as the stages before it. The participants were again asked to pair compare each of the sets choosing the one they find as more realistic. They were informed about the total number of comparison (100) and told that they have a total of 30 minutes to complete the test (20 seconds for an image). The displacement shift sets were made of 100 pairs including all of the comparisons in between the 25 pairs 2 times. This is as for 5 images per 5 views and a comparison in between just the ones from the same view the total comparisons are 50.

8.4. Second Experiment Conclusion

The conclusions of the two stages are compared according to L.L.Thurstone's Law of comparative judgment (20) and scaled by the mean number of choices that each user gave to the different samples. The second methodology (counting choices) allows for a better overview on the specifics of the preference of the participants involved. It is generated in Excel while the first in Matlab utilizing a statistic toolkit library. The second is presented for both the angle shift and the placement distribution experiments (Fig. 22 and 23). These present the mean choices converted in percentage according to the choices of the participants. What is immediately noticeable is the small differences the samples excite for the perception of realism in between them. In the first graph (Fig.22) only the sample with no angular shift scores relatively low compared to the other two samples. The difference in between the other two is negligible. In the second graph (Fig.23) all of the samples including the one which represents the base state (10_0) are evenly distributed. The angles shift sample added to the displacement (5_5) scores the lowest in the comparison.

0 cm. of displacement Fig.19

5 cm. of displacement Fig.20

10 cm. of displacement Fig.21

Thurstonian scaling (20) was also used to analyze the data. For every level within one attribute a Z-score was calculated, and the level which was preferred the least got a Z-score of 0. The delta Z-score could then be easily interpreted and calculated into percentages that one attribute is preferred over the other. As a general indication, a delta Z-score of 0, 0.5, 1, and 1.5 corresponds respectfully with the percentages 50%, 70%, 85% and 93%. For example, when the first level has Z-score of 0, 5 and the second a Z–score of 1.0, the delta Z-score is 0.5. This means that 70% of the people preferred the second level over the first.

Fig. 22

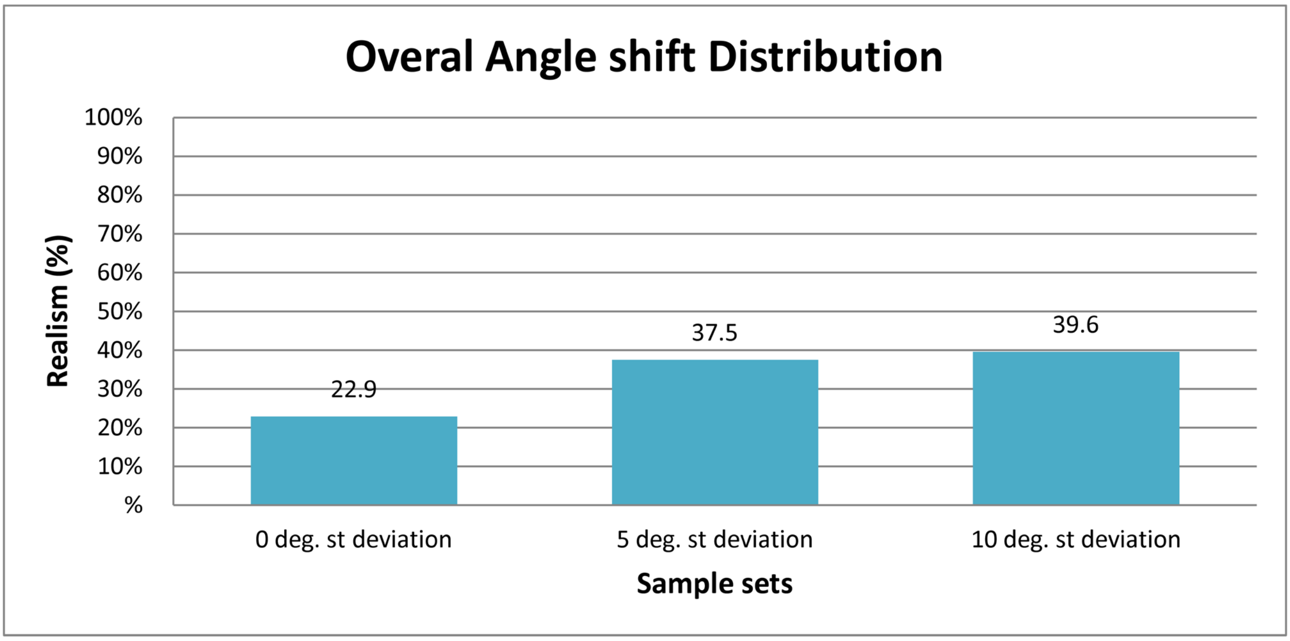
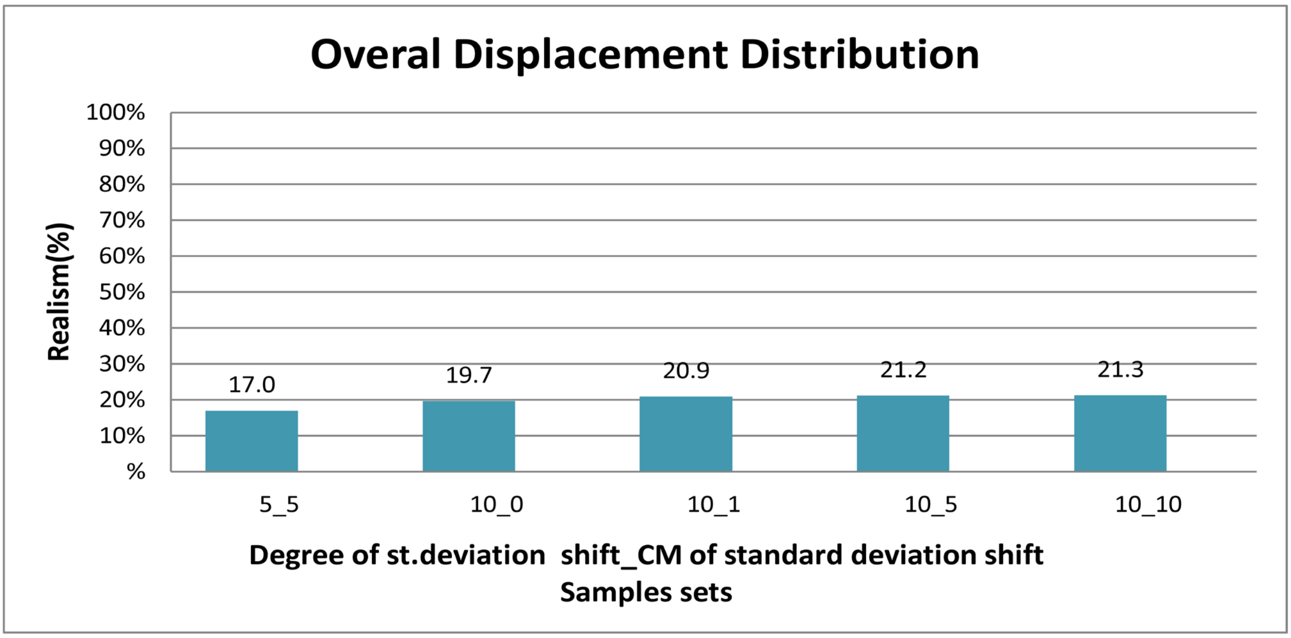
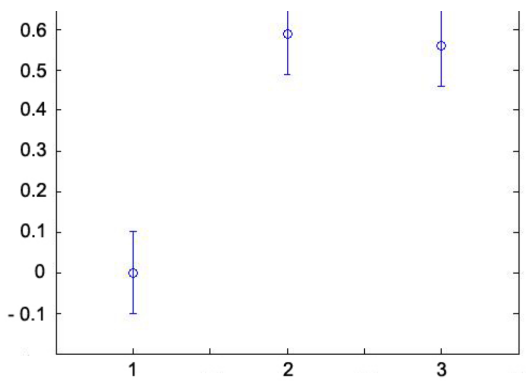


Fig. 23



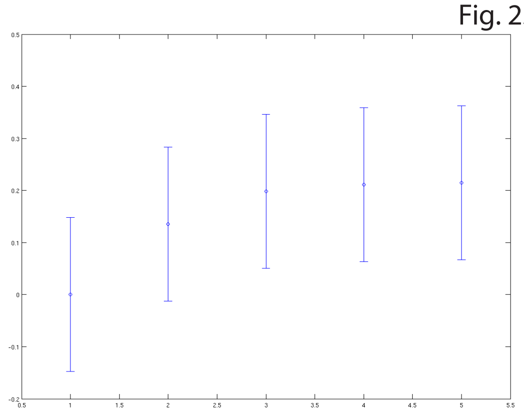
To make it possible to detect statistical significance when interpreting graphs, error bars of this method were created. To calculate these error bars for each level we used Montags' method (21). Therefore, a significant difference ($p>0.05$) is present when the items are outside of the error bars of their comparisons. This means that for specific samples the area standing outside of the error bars translates to a 95% confidence that the outcome has a statistical significance. According to this for the rest of this project we only mention comparisons and analyze samples standing outside of each others' error bars.

Fig. 24



Chance of choice for angle shift. (1 = 0 deg.; 2 = 5 deg.; 3 = 10 deg.;)

Fig. 25



Chance of choice for displacement. (1 = 5 deg. 5 cm.; 2 = 1 deg. 0 cm.; 3 = 10 deg. 1 cm.; 4 = 10 deg. 5 cm.; 5 = 10 deg. 10 cm.;)

It's clearly seen that the statistical probability that one sample set would be chosen over the other is not significant in Fig. 21. The greatest difference is 57 % and it stands between sample set 1 and sample set 5. They are also the only two standing outside of each others error confidence bars for this analysis. For the angle shift (Fig. 22) however, we can see that samples number 2 and 3 are 72 % more likely to be chosen than sample number one. Still the confidence intervals show that 3 is not more likely to be chosen over 2 and vice versa.

The data for both of the sets is also analyzed according to the view in which the participants made the pair choice. This is for both the angle shift (Fig.26) and the displacement (Fig.27). As you can see the trendlines for both show a very small difference in the effect of the view on the judgment of the participants. This combined with the overall small differences in between the realism of different samples shows that the images excite a change in the perception of realism which is of very small significance.

For the final stage of the experiment we take a deeper look into the 0 deg. clutter samples which have the greatest effect on the realism and score considerably lower that the other they are pair compared with.

Fig. 26

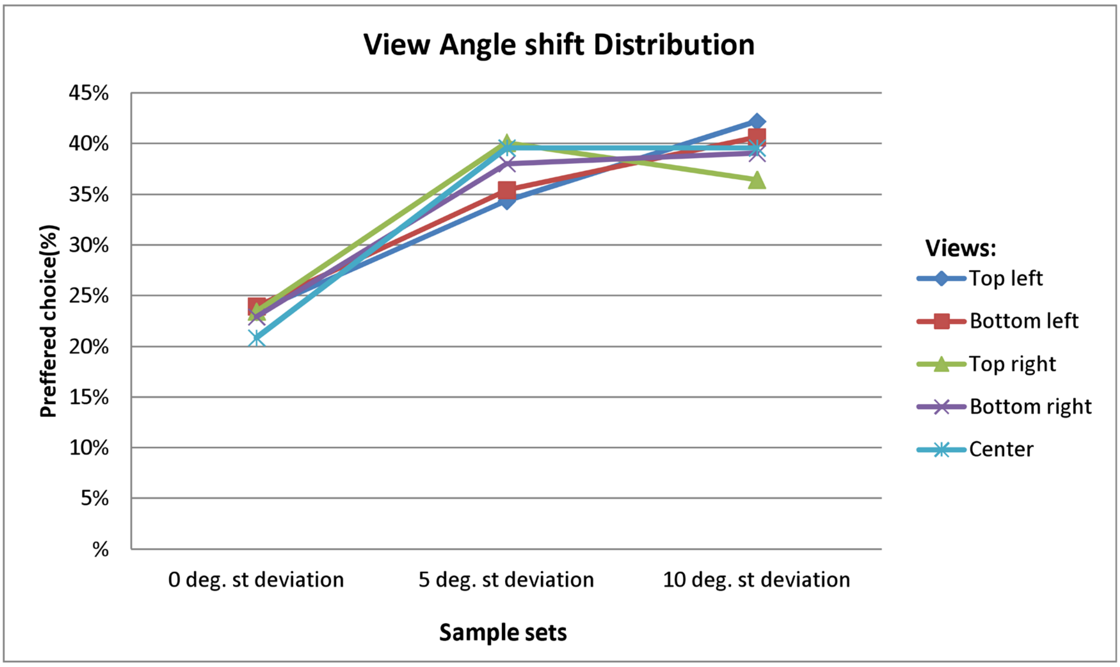
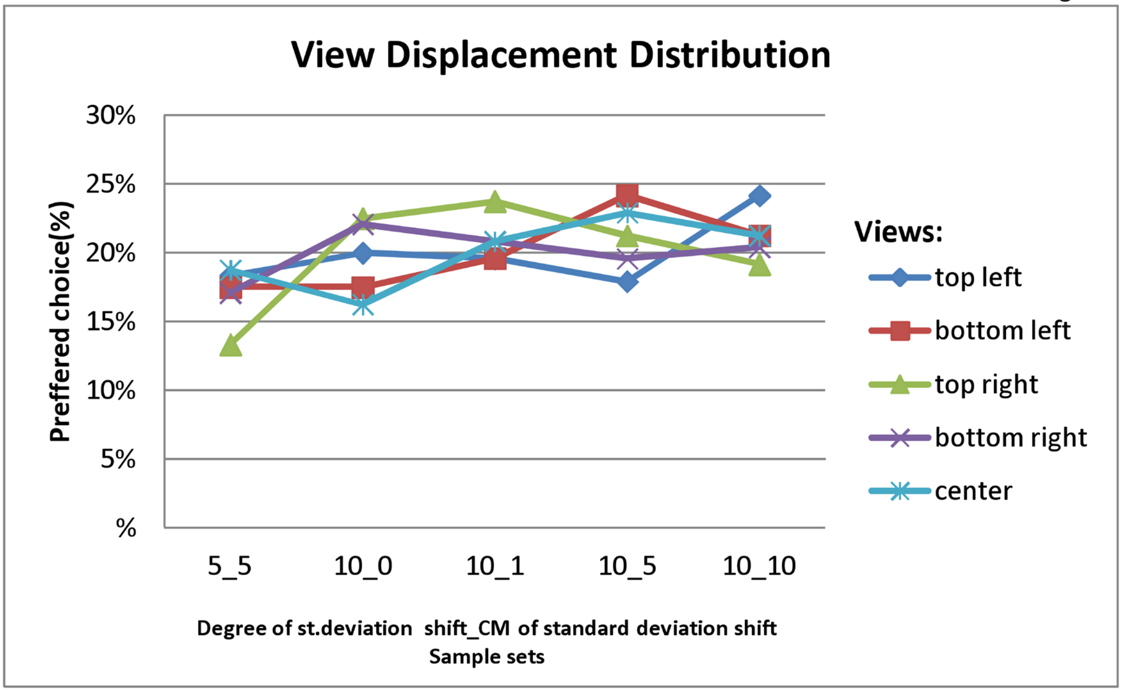

View Angle shift Distribution
45%
40%
35%
30%
25%
20%
15%
10%
5%
%
Preffered choice(%)
0 deg. st deviation
5 deg. st deviation
10 deg. st deviation
Sample sets
Views:
Top left
Bottom left
Top right
Bottom right
Center

Fig. 27


View Displacement Distribution
30%
25%
20%
15%
10%
5%
%
Preffered choice(%)
5_5
10_0
10_1
10_5
10_10
Degree of st.deviation shift_CM of standard deviation shift
Sample sets
Views:
top left
bottom left
top right
bottom right
center

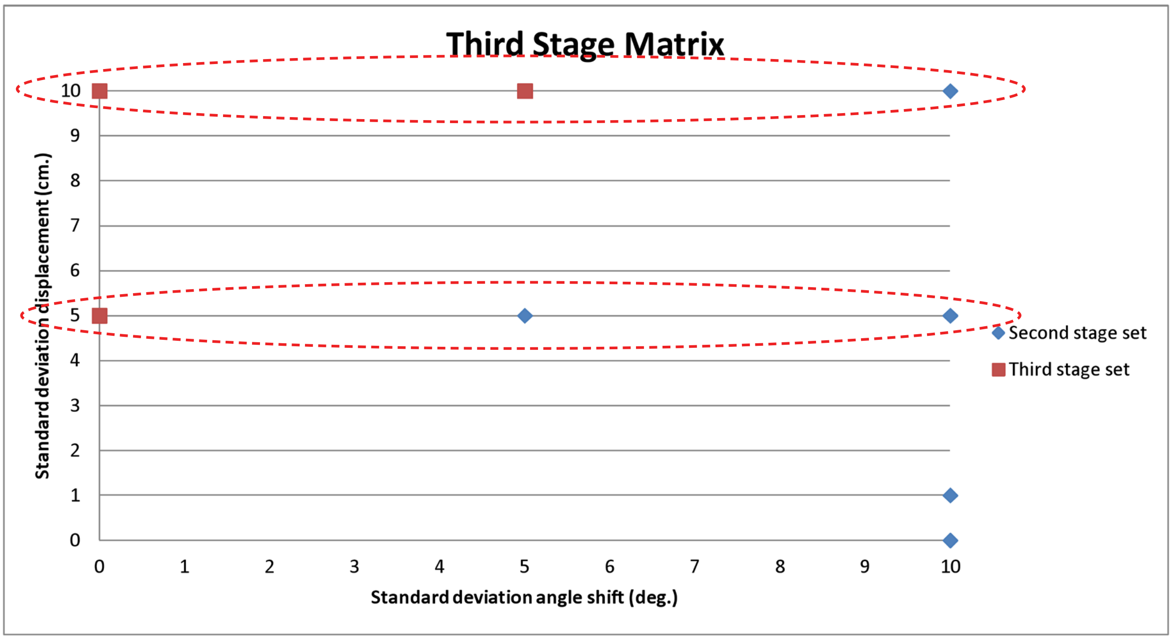
9. Third Experiment

For the third user test an additional level of comparisons is used. This final stage evolves around the outcome of the 0 deg. test sample in the standard deviation angle shift analysis (Fig.22). We found this sample to be particularly interesting as the one affecting the perception of realism the most. Thus for the third experiment we found it necessary to create a matrix covering some of the samples already created and adding ones which will complete the sets needed to create the matrix. This matrix gives us more quantitative data on the outcome of the sample of interests. In this stage the scene and the methodology are the same as the previous second test. The difference in the variables however can most easily be understood when a look at the actual matrix is taken (Fig.28). What can be seen in it, is the position (according to angle and displacement) that the samples available and the samples created have. The third stage set necessary for the experiment are the sets at 5 cm. and 10 cm. of standard deviation displacement (Fig.28).

9.1. Test Methodology

Six out of eight shown samples (outlined) are compared between each other The two groups are pair compared in a matrix thus each angle and displacement set is compared with every other while again there are no comparisons in between different views. For all the participants this totals to 75 comparisons (15 comparison per view x 5 views). For this stage the participants are again given 30 minutes to evaluate which of the two shown samples they perceive as more realistic than the other.

Fig. 28



9.2. Third Experience Conclusions

The conclusions of the third user test are presented in the same way as the conclusions from the second. Figure 29 represents the mean choice percent for each of the sample sets in respect to their realism. What's immediately visible is that there is a small rise in realism with the rise in displacement and angular spread. In figures 30, 31, 32 you can view the effect of the angle shift on realism with respect to the specific comparisons, while in Fig. 33 the effect of the displacement is shown. For the displacement this analysis produces a mean equal preference for the two steps (5 cm and 10 cm) when compared to their equivalent in all the angle shifts steps (Fig.33). In Fig. 34 the cross effect can be seen.

All the graphs represent the percentage of times in which one image was chosen in respect to its realism over the other in the specific comparisons. As said before the choice was made to only make the pairwise comparisons while not mixing between views. It was previously demonstrated that the views have no significant effect on the outcomes. In the following graphs (Fig 35,36,37) all of the analysis are redone with the modified Montag method allowing for confidence intervals of 95% to be set for each of the shown comparisons. These graphs again demonstrate the statistical significance of the choices in respect to the change (in percentage) that one images is chosen over the other while compared in pairs.

Fig. 29

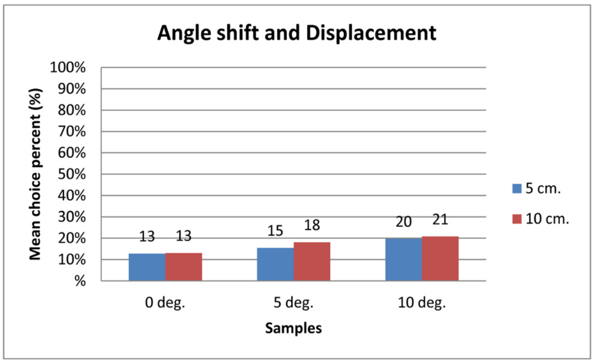


Fig. 30

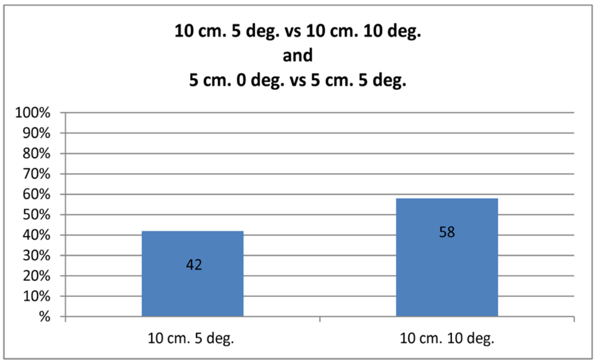


Fig. 31

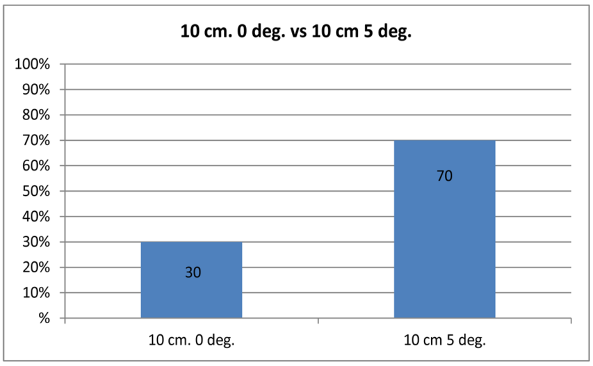


Fig. 32

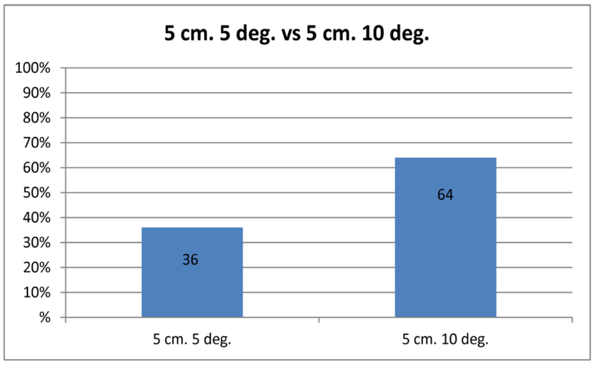


Fig. 33

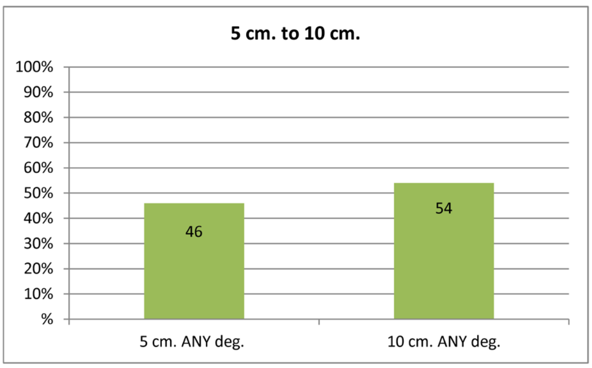


Fig. 34

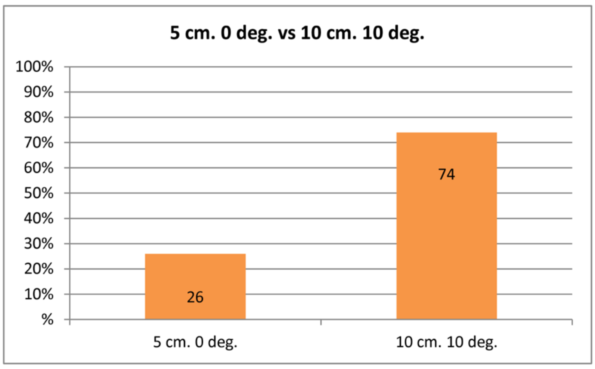
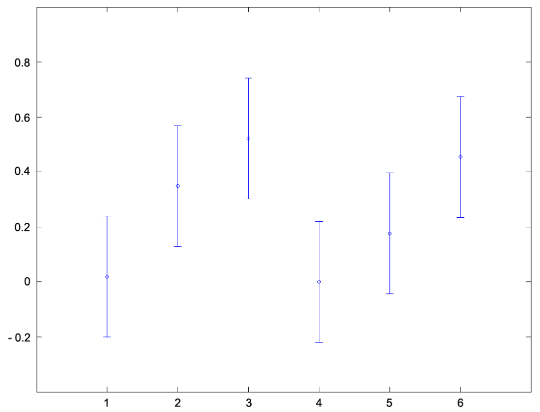
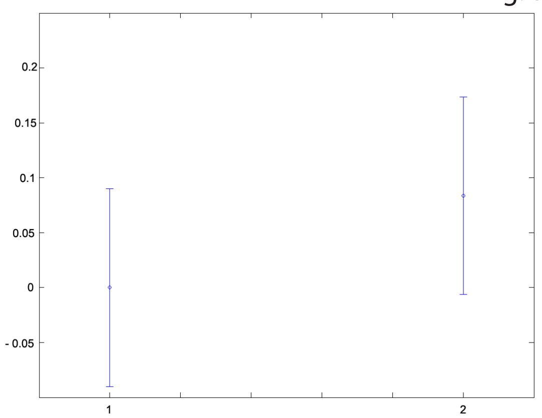


Fig. 35



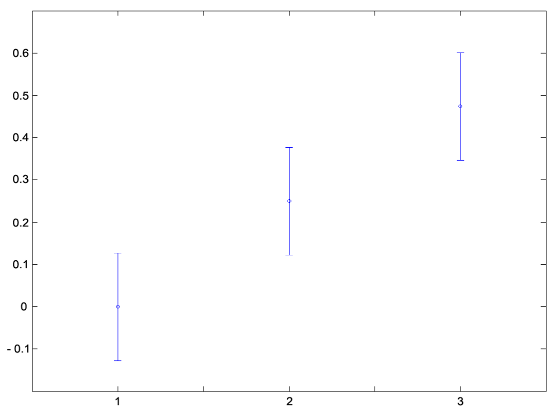
Chance of choice for all samples. (1 = 10 cm 0 deg.; 2 = 10 cm 5 deg.; 3 = 10 cm 10 deg.; 4 = 5 cm 0 deg.; 5 = 5 cm 5 deg.; 6 = 5 cm 10 deg.;)

Fig. 36



Chance of choice between grouped samples for 5 cm. and 10 cm. (1 = 5 cm ; 2 = 10 cm.)

Fig. 37



Chance of choice between grouped samples for 0 ,5 and 10 deg. (1 = 0 deg; 2 = 5 deg; 3 = 10 deg;)

What can be seen in Fig. 35 is the relation between all of the sample sets used in the paired comparisons. The graph corresponds to Fig. 29. What is immediately noticeable is that there is no statistical significance between the 5 and the 10 cm. samples in relation to their realism. However that is not the case with the changes in angle shifts. There is a small but steady rise in realism with the rise in angle shift. What can also be seen here is that the cross effect (given in Fig.34) is statistically significant and there is a 68% chance that sample number 6 will be chosen over sample number 1.

In Fig. 36 all samples for the 5 cm. and the 10 cm. of standard deviation are grouped and compared. As you can see, there is no statistical significance. This graph corresponds to the graph in Fig. 33. Given the fact that the percentage of rise is very small and the confidence intervals show no significance between the samples it can immediately be seen that the displacement again did not affect the perception of realism with any statistical significance.

Fig. 37 shows the summarized samples for the angle shift levels. These are (refer to Fig.29) the samples in the 0 , 5 and 10 deg. sets. In this graph it can be seen that the percentage differences in figures 30, 31, 32 are significant. The chance of choice between the samples are roughly 59% for a one level rise and 68% for two.

10. Project Conclusions

On the basis of the experiment made we can conclude that under the specific conditions set in the experiments clutter did not affect realism greatly. Therefore the single answer to the research question of this project is that:

" The perception of realism is not greatly affected by the amount of clutter in the scene"

Despite the fact that there is a proven small and consistent rise in realism with the rise of the clutter that rise was not affected by either the number of objects or their placement in the scene but only by the variable of angle spread. As we demonstrated in the second and third experiment, under the conditions set in this project angle shift proved to have a greater effect on the visual perception of realism than the other features of clutter. This was consistent over the samples of the last stage, which we believe is the most detail revealing for the entire project. Still the outcome of the second angle experiment showed that there is no difference in perceived realism between the 5 deg. and 10 deg. samples (Fig.12). This leads us to believe that the perceptual value the angle shift to the visual realism of scenes could be even smaller in a none test environment. Thus the above stated research answer is given. We believe that the methodologies as well as the high number of images used in this project stands as additional verification for the validity of the derived data and subsequently for the accuracy of the above stated research answer.

As this study was designed to provide the researchers with specific quantitative analysis, the additional talks made after each stage have not been mentioned so far. However in an informal discussions after all the experiments users pointed mostly towards objects and their material and surface features for what they thought affects the realism of the scene the greatest. These were ranging from material and color specifics to shadow softness preferences. However only a few of all the participants pointed towards either one of the variables of clutter as a factor affecting their judgment of which images look more realistic. Thus while in specific cases there is some rise in realism with the rise in clutter it definitely needs to be specified that this might be a reason to look into other features of virtual scenes and see what their affect on the perception of realism is. According to the qualitative data derived in the informal discussions we might also have the predisposition that some of these other features of the environments, which are not deplorable to the constant overall render quality, might have a greater effect on the perception of realism than clutter. However in order to say that with certainty it needs to be scientifically proven.

For professional luminaire designers we believe the current project can serve as a demonstration to what the actual perceptual value of clutter is, while still minding the fact that the resources spent in projects are in most instances case specific. For virtual prototyping of luminaires the outcome of this project points towards a defened simplification of the scenes used. As the number of object in the scene does not greatly contribute to the perceptual value of the renderings the 3d scenes can be stripped off objects inside them and thus made quicker, with less skill required and eventually rendered faster. However we stay far from trying to stipulate that adding any objects to 3d scenes is pointless. This is as besides the actual perceptual value some objects in 3d environments obviously also carry other values such as for instance the artistic. These are yet to be quantified. Nevertheless it is arguable if these are of importance to luminaire prototyping. Still, if we believe that this is the case, it is also questionable what objects should be chosen to be most informative for the characteristics of the luminaires and if not chosen by this criteria then in what way are they contributive to the rendered environments and virtual luminaire prototyping ?

11. Recommendations

There are many directions which virtual 3d prototyping of luminaires can take. These more generally can be divided on account of the resources spent and the application of the renderings.

11.1. Specialized Applications

For future advances in specialized luminaire prototyping further research needs to be done which looks at other specific features of virtual environments. It is also important to eventually quantify the perceptual value (as connected to realism) that renderings have when compared to the ground truth. These should be done with either commercial or specialized software packages when all the specific methods are carefully benchmarked between one other so that the technique and software package that produces the most visually realistic results can be specified. Still we believe that a 3d designer which is skilled with a set of industry specific software could be able to compensate for the differences in realism given the fact that it is not expected that for specialized cases the same environment set up is used over a couple of projects. Thus it is important to confirm any outcomes with professionals involved in creating the industry standards. This is also as so far, the amount of human resources needed for utilizing different methods is not quantified and only seasoned users can be referenced for specifics about the time and skill needed to utilize different techniques. Nevertheless we can claim that, at the current moment, specialized virtual luminaire prototyping is successfully done as it is largely dependent on the clients' personal preferences. These however are also affected by numerous known and unknown factors. Still we believe it is entirely up to the industry professionals to constantly push the quality standards which (we believe) are now arguably constrained by either software or hardware technologies.

11.1. Commercial Applications

The implications for research and development in the field of commercial virtual luminaire prototyping are also numerous. These future software packages can be either online or offline applications developed with comprehensive interfaces which allow customers to render their preferred objects or environments illuminated by the luminaires which they have chosen and can subsequently acquire. The downfall in this is that currently the no approximation rendering techniques are too demanding in terms of computations and more planning is needed for them to be developed into commercial user packages. However the outcome of this project shows that the scenes used in such, can be simplified to a great extend while still remaining visually realistic. Therefore the development of such systems and their application in smart phones, tablets and other low end hardware systems is imminent.

Nevertheless the current developments in image based rendering including packages incorporating light simulation (22,23) can already be compared with the ground truth. This is while other work focused on extracting light reflectance fields from photography (24,25) proves that new techniques can lead to improvement of the already existing packages. However for the current diversity in luminaires that a company such as Philips provides it is highly improbable that any other format than photometric data files (IES, LDT) can be immediately utilized. Adding this to a list of obstacles that a virtual luminaire prototyping system has to overcome, makes it highly probable that a versatile user focused software needs to be developed.

References:

1. Wikipedia the free online encyclopedia; Light; http://en.wikipedia.org/wiki/Light

2. Realism in Computer Graphics; Andries van Dam; Brown University October 2011;

3. 3d Photorealism Toolkit; Bill Fleming 1998;

4. The free dictionary by Farlex; The definition of “clutter” (http://www.thefreedictionary.com/ clutter);

5. Psychophysically based Artistic Techniques for Increased Perceived Realism of Virtual Environments; Peter Longhurst, Patrick Ledda, Alan Chalmers; University of Bristol, Merchant Ventures Building, Woodland Rd, Bristol, BS8 1UB;

6. A comparison of perceived lighting characteristics in simulations versus real-life setup; B. Salters*, P. Seuntiens Philips Research Eindhoven, High Tech Campus 34, Eindhoven, The Netherlands;

7. Wikipedia the free online encyclopedia; Bidirectional reflectance distribution function; (www.wikipedia.org/wiki/Bidirectional_reflectance_distribution_function);

8. The Convergence of Graphics and Vision; Jed Lengyel; In IEEE Computer, July 1998;

9. Measuring the Perception of Visual Realism in Images; Paul Rademacher, Jed Lengyel Edward Cutrell, Turner Whitted, University of North Carolina at Chapel Hill and Microsoft Research;

10. Wikipedia the free online encyclopedia; Tone mapping; (http://en.wikipedia.org/wiki/Tone_mapping);

11. Wikipedia the free online encyclopedia; 3ds Studio MAX; (www.en.wikipedia.org/wiki/Autodesk_3ds_Max);

12. Wikipedia the free online encyclopedia; Indigo Renderer; (www.en.wikipedia.org/wiki/Indigo_renderer);

13. From Photon to Photography; Indigo Renderer; (www.indigorenderer.com);

14. Luminance 2.20 HDR; Freeware HDR Image Processing Software; (www.qtpfsgui.sourceforge.net);

15. Photographic Tone Reproduction for Digital Images; Erik Reinhard Michael Stark Peter Shirley James Ferwerda; University of Utah;

16. Display Calibrations Explained; Understanding Delta-E; (www.displaycalibrationonline.com/colorscience_delta.asp);

17. The choice between a five-point and a ten-point scale in the framework of customer satisfaction measurement; Pedro S. Coelho, Susana P. Esteves New University of Lisbon;

18. Wikipedia the free online encyclopedia; L.L.Thurstone – Law of comparative Judgment; (www.wikipedia.org/wiki/Law_of_comparative_judgment);

19. Psychometric Scaling: A toolkit for imaging Systems Development; Peter G.Engeldrum; Winchester, Massachusetts, USA; year 2000;

20. Wikipedia the free online encyclopedia; L.L.Thurstone - Thurstone Scale; (www.wikipedia.org/wiki/Thurstone_scale);

21. Empirical formula for creating error bars for the method of paired comparison; Ethan D. Montag; Rochester Institute of Technology;

22. Rendering Synthetic Objects into Legacy Photographs; K.Karsch, V.Hedau, D.Forsyth, D.Hoiem; University of Illinois at Urbana-Champaign;

23. Acquiring the Reflectance Field of a Human Face; P.Debevec, T.Hawkins, C.Tchou; University of California at Berkeley and LifeF/X, Inc;

24. Image-based Separation of Diffuse and Specular Reflections using Environmental Structured Illumination; B.Lamond, P.Peers, A.Ghosh, P.Debevec; University of Southern California, Institute for Creative Technologies;

25. Inverse Global Illumination: Recovering Reflectance Models of Real Scenes from Photographs; Yizhou Yu, P.Debevec, J.Malik, T.Hawkins; Computer Science Division, University of California at Berkeley;

TU Delft